

Natural Selection at the NHLH2 Core Promoter Exceptionally Long CA-Repeat in Human and Disease-Only Genotypes in Late-Onset Neurocognitive Disorder

Hossein Afshar^a Fatemeh Adelirad^b Ali Kowsari^c Naser Kalhor^c
Ahmad Delbari^a Reza Najafipour^d Mahshid Foroughan^a Ali Bozorgmehr^e
Safoura Khamse^a Neda Nazariapanah^b Mina Ohadi^a

^aIranian Research Center on Aging, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran;

^bDepartment of Health Education and Promotion, Faculty of Health Sciences Tabriz University of Medical Sciences,

Tabriz, Iran; ^cDepartment of Mesenchymal Stem Cell, The Academic Center for Education, Culture and Research,

Qom, Iran; ^dCellular and Molecular Research Centre, Research Institute for Prevention of Non Communicable

Disease, Qazvin University of Medical Sciences, Qazvin, Iran; ^eDepartment of Neuroscience, Faculty of Advanced Technologies in Medicine, Iran University of Medical Sciences, Tehran, Iran

Keywords

NHLH2 · Short tandem repeat · Neurocognitive disorder · Natural selection

Abstract

Background: Approximately 2% of the human core promoter short tandem repeats (STRs) reach lengths of ≥ 6 repeats, which may in part be a result of adaptive evolutionary processes and natural selection. A single-exon transcript of the human nescient helix loop helix 2 (NHLH2) gene is flanked by the longest CA-repeat detected in a human protein-coding gene core promoter (Ensembl transcript ID: ENST00000369506.1). NHLH2 is involved in several biological and pathological pathways, such as motivated exercise, obesity, and diabetes. **Methods:** The allele and genotype distribution of the NHLH2 CA-repeat were investigated by sequencing in 655 Iranian subjects, consisting of late-onset

neurocognitive disorder (NCD) as a clinical entity ($n = 290$) and matched controls ($n = 365$). The evolutionary trend of the CA-repeat was also studied across vertebrates. **Results:** The allele range was between 9 and 25 repeats in the NCD cases, and 12 and 24 repeats in the controls. At the frequency of 0.56, the 21-repeat allele was the predominant allele in the controls. While the 21-repeat was also the predominant allele in the NCD patients, we detected significant decline of the frequency ($p < 0.0001$) and homozygosity ($p < 0.006$) of this allele in this group. Furthermore, 12 genotypes were detected across 16 patients (5.5% of the entire NCD sample) and not in the controls (disease-only genotypes; $p < 0.0003$), consisting of at least one extreme allele. The extreme alleles were at 9, 12, 13, 18, and 19 repeats (extreme short end), and 23, 24, and 25 repeats (extreme long end), and their frequencies ranged between 0.001 and 0.04. The frequency of the 21-repeat allele significantly dropped to 0.09 in the disease-only genotype compartment ($p <$

© 2020 S. Karger AG, Basel